Alternative color filter array layouts for digital photography

K. Gorokhovskiy, J.A. Flint and S. Datta

Department of Electronic and Electrical Engineering,
Loughborough University, Loughborough, LE11 3TU, UK
kgorokhovsky@apical-imaging.com
J.A.Flint@lboro.ac.uk, S.Datta@lboro.ac.uk

Abstract— **The performance of digital cameras depends not only on the accuracy of methods of restoration of missed color samples (demosaicking) for a given color filter array, but also from spatial configuration of color sensors in the color filter array (CFA) itself. This paper considers three different color filter array (CFA) patterns; the established (2 by 2) Bayer pattern, the 3 by 2 (6-sample) and the 3 by 3 diagonal Bayer CFA. One difficulty in comparing the different schemes is the influence of the demosaicking algorithm on the result. In order to remove this dependence we propose three methods of comparison. They are (a) measuring widowed averages of colors on large areas (b) visual comparison of interference between regular patterns of images and CFA, and (c) utilization of one layer neural networks to build demosaicking algorithm for selected color filter arrays. A substantial image database comprising 1338 images has been used to experimentally validate the different patterns.**

I. INTRODUCTION

Since the introduction of the Bayer color filter array [1] in 1975 digital photography has seen dramatic progress. This progress has converted digital photography from an exclusive tool available only to the aerospace industry to widespread everyday use in consumer digital photo cameras. One feature of digital photography which has remained almost unchanged throughout the decades is the layout of color components in the photo-sensitive array. The de-facto standard is the Bayer color filter array which is proven to be technologically feasible and robust enough for wide variety of applications. Although alternatives are available on the market (such as Foveon CFA) their widespread uptake has been suppressed due to technological difficulties in production which lead to increased costs of the produced electronic chips. The Bayer color array was based on knowledge of human visual perception available at that time and is consistent with the theory of YUV color space. The purpose of this paper is to evaluate an optimal distribution of color components in a color filter array where the YUV model is not assumed. In this paper several alternative color

patterns will be compared and appropriate performance metrics will be introduced.

II. SELECTING COLOR PATTERNS FOR EVALUATION

Since there are an infinite number of possible color filter layouts it is necessary to select a small number of candidates for further evaluation. The selection of suitable patterns from all the possible configurations is a complex task which is not described in detail here, however the constraints used are (briefly):

- Minimization of the atomic building block for pattern. In other words it could be a 2x2, 3x2 or 3x3 pattern.

- The frequency of color components should be similar in horizontal and vertical dimensions

- The frequency of color samples should be consistent with an established theory of human visual perceptional sensitivity [2].

After removing trivial cases the following patterns were selected for evaluation:

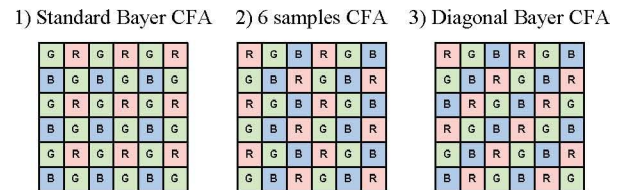1) Standard Bayer CFA   2) 6 samples CFA   3) Diagonal Bayer CFA

Figure 1.   Spatial layout of patterns used for evaluation

The left hand layout pattern in Figure 1. represents the classic Bayer pattern which is widely used nowadays in wide range of digital cameras. The building block in the Bayer pattern is 2 by 2. The central layout represents next possible size of building block (3 by 2). On the contrary to classic Bayer CFA the colors are mixed here in equal proportions. One disadvantage of the 6 sample CFA is that it is not symmetrical to permutation of vertical and horizontal dimensions. This type asymmetry is removed in diagonal Bayer CFA (right hand pattern in Figure 1). It has a building block with dimensions 3 by 3 and in common with the 6-sample CFA, has an equal number of each color filter element.

III. METHODS OF COMPARISON

Three different methods were selected to compare these CFAs:

- Calculation of the average value of color for a local area to compare mosaicked and original images on an image database
- Evaluation of errors in the average colors on synthetic patterns
- Creation of demosaicking algorithms using neural networks to evaluate their performance

Many modern robust demosaicking algorithms use implicit statistical assumptions about possible configurations of small image details. Although they are mostly heuristic they work well for most cases. However, the exceptions are high frequency repetitive patterns where it is impossible to restore all data due to Nyquist limit. In this case the algorithms can misinterpret not only local details but also the mean colors of large areas. The ability to restore mean colors is highly dependent on configuration of CFA layout.

A. Evaluating errors in color on a large image database

Due to interference between image repetitive structures and the repetitive pattern of color filer arrays, errors in interpolating color components can accumulate even on areas which are significantly bigger than the demosaicking pattern itself. Using a representative image database it is possible to determine which pattern minimizes the color averaging artifacts on areas bigger than minimal repetitive block of color filter arrays.

The UCID database version 2 from [3] was used to perform the measurements. It contains 1338 uncompressed images. Each image has a size of 512×384. The images were captured using a Minolta DiMAGE 5 camera and stored in uncompressed TIFF format. The original size of CCD matrix of Minolta DiMAGE 5 camera is much larger than the final image size and therefore high frequency details are well presented. Thus, it is assumed that the reference images are relatively unaffected by the quality of optical system or the demosaicking algorithm used in this camera. The original images were mosaicked again by selective sampling of the RGB data using the patterns given in Figure 1 for the experiment. Then, the following formula has been used to calculate the average colors for given pixel:

$$\overline{R} = \frac{1}{c_R} \cdot \sum_y \sum_x R_{xy} \cdot M_{xy}^R \cdot e^{-\frac{x^2+y^2}{\gamma}} \tag{1}$$

$$c_R = \sum_y \sum_x M_{xy}^R \cdot e^{-\frac{x^2+y^2}{\gamma}} \tag{2}$$

These formulas were used for calculation of red average, $\overline{R}$, however similar expressions were used for the other color planes. M_{xy} are matrices for demosaicking composed of zeros and ones. A value of 1 in M_{xy} represents the presence of a given color detector at position (x, y) and a value of 0 its

absence. Equation 1 has been also used to calculate the average value for color components. To do this, all elements of matrix M were assumed to be unity. In other words we have all color detectors in every position (x, y). For the purpose of quantifying the errors introduced by the array patterns, the mean square error, MSE has been calculated. However, it is recognized that this metric does not account for the differences that a human would perceive between original and restored pictures. To estimate the visual difference the following simple model was used [2], [4] and [5]. The areas with a small deviation in color are not distinguishable for human eye from those parts where the color is exact. There is some threshold beyond which the difference in color becomes visible. The theory suggests that for every combination of R, G and B components there exists a three dimensional ellipsoid of perceptually equivalent changes.

$$\begin{cases} |R_o - R_r| \le 0.02 \\ |G_o - G_r| \le 0.02 \\ |B_o - B_r| \le 0.04 \end{cases} \tag{3}$$

Where $R_o, G_o, B_o \in [0,1]$ are original colors and $R_r, G_r, B_r \in [0,1]$ are the restored colors. For our simplified model the formula (3) gives a volume in color space where samples are indistinguishable for the human eye. The volume is greater than the average volume of distinguishable colors. The capability to resolve colors by a human will also depend on the output device. We have chosen the maximum volume in order to detect the worst case, i.e. when errors in color restoration would be visible on almost any output device.

Another metric used in the evaluation was Normalized Color Difference (NCD). NCD quantifies the perceptual color difference and is defined as follows:

$$NCD = \frac{\sum_{x,y} \sqrt{\left(L_o^{(x,y)} - L_r^{(x,y)}\right)^2 + \left(U_o^{(x,y)} - U_r^{(x,y)}\right)^2 + \left(V_o^{(x,y)} - V_r^{(x,y)}\right)^2}}{\sum_{x,y} \sqrt{L_o^{(x,y)2} + U_o^{(x,y)2} + V_o^{(x,y)2}}} \tag{4}$$

Where L, U, V are lightness and chrominance components of the resultant and original images in CIELUV color space. They are converted from image sRGB samples through CIEXYZ color space. The conversion sequence was the following: sRGB to linear RGB to CIEXYZ to CIELUV.

In total, 1338 images from the database were used to compute MSE and NCD [6] of color components and the number of samples which exceed the bounds of the volume defined in Equation (3) were found. A standard deviation figure was also calculated for both these characteristics. As can be seen from Table 1, the diagonal Bayer CFA outperforms the standard Bayer pattern with the used color perception model. The 6 sample pattern is

266

generally inferior to the standard Bayer CFA. The errors produced by it are greater then for other two arrays.

TABLE I.   ERRORS OF AVERAGE COLOR

	Standard Bayer 2x2	6 samples CFA	Diagonal Bayer 3x3
NCD	0.0183	0.0185	0.0177
R channel MSE	0.0123	0.0127	0.0114
its deviation	0.0053	0.0058	0.0050
G channel MSE	0.0108	0.0119	0.0108
its deviation	0.0047	0.0052	0.0047
B channel MSE	0.0111	0.0109	0.0103
its deviation	0.0047	0.0047	0.0045
average number of visible differences	314.39	361.50	277.63
deviation of number of visible differences	153.25	173.17	139.35

B. Visualization of errors on synthetic images

As has been explained, the demosaicking of periodic structures can be the most challenging cases. The CFA has a tendency to interfere with these producing visible artifacts and as a result, false colors appear. In most cases the periodic structures have an artificial origin such as fences, buildings, and so on. In general, any discrete repetitive structure with size of repetitive block $N \times M$ can be written in the form of a Fourier-series as:

$$I(x,y) = \sum_n \sum_m C_{nm} \cdot e^{i(xn+ym)} \qquad (5)$$

Where I is the pixel intensity. The structures which can be generated vary significantly and can produce any combination of pixels in a block. This makes exploration of interference more difficult as it requires more computations. To minimize experiment computation time only one-dimensional waves were used. They can be represented using the following subspace:

$$I(x,y) = e^{i(A \cdot x + B \cdot y + C)} \qquad (6)$$

As will be shown below this simplification allows easy visualization on two dimensional images. In addition this kind of structures is common in images.

The procedure of visualization can be described as follows: for every a and b with given N and M find maximum difference between windowed average of original image and windowed average of restored (from mosaicked) image.

$$R(a,b) = \max_{c \in [0,2\pi]} \left| \sum_{x=-\frac{N}{2}}^{\frac{N}{2}} \sum_{y=-\frac{M}{2}}^{\frac{M}{2}} \left(1 - M_{xy}^R\right) \cdot e^{i(a \cdot x + b \cdot y + c)} \cdot e^{-\frac{x^2+y^2}{\gamma}} \right| \qquad (7)$$

Where M_{xy}^R describes coefficients of mosaicking pattern for the red channel. The formulas for green and blue channels are similar. The resultant $R(a,b)$, $G(a, b)$, and $B(a,b)$ can be joined together and viewed as a color image. Results of the visualization for the CFAs being considered are shown in Figure 2 below. The dark areas represent local error maxima; brighter areas show smaller errors. The centre of each image represents the condition $a = 0$, $b = 0$.
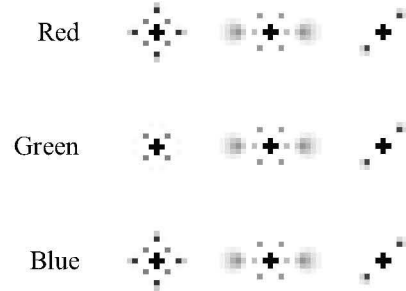

Figure 2.   Results of visualization of artifacts for Bayer CFA (left), 3×2 CFA (center), Diagonal Bayer 3×3 CFA (right)

The images illustrate clearly which waves are most difficult for interpolation for given filter and color channel. Only one pattern is asymmetrical with respect to colors. It is classic Bayer array. As can be seen, the green channel has very low interference but red and blue channels are sensitive to horizontal and vertical waves with wavelength of $4 \cdot \Delta$, where Δ is distance between centers of two horizontal or vertical cells. Second and third layouts taken for comparison demonstrate equivalent configuration of errors for every color plane. This is logical as they are symmetric with respect to the permutation of colors.

As can be seen, the most sensitive to low frequency waves is the 6 pixels CFA. Then, for the Bayer CFA errors for low frequencies exist for red and blue channels only (4 green pixels). The diagonal Bayer CFA is ineffective on diagonal waves but it is efficient in both horizontal and vertical directions. The visualization does not provide a quantitative measure of errors in color interpolation; however it demonstrates clearly the weaknesses in the imaging of repetitive structures for each CFA.

C. Utilization of neural networks for comparison of color arrays

The algorithms used in demosaicking differ depending on the configuration of the CFA. As a result it is difficult to compare the pure performances of different layouts of color filter arrays for demosaicking. Back-propagation linear neural networks can be used to minimize this dependence. Neural nets were selected because it is possible to build a demosaicking algorithm using a training process. The structure of the nets can be the same for all layouts. Only biases will differ. This can possibly give us an assurance that the differences of performance of demosaicking will be not a result of robustness of an algorithm but mainly result of

layout of a pattern. Although, use of non-linear neural networks can probably show better performance, simple linear back-propagation neural nets have been chosen as they provide:

- At least one and only one optimum for given training set and a neural net configuration

- Due to their simplicity they provide faster learning

- The theory of linear methods is already well developed as shown in [7]

Alternatively it is possible to use linear regression methods but it was found that the tools available for neural networks provide more control on the process of finding an optimum. The neural network configuration used in the experiments is depicted in Figure 3.
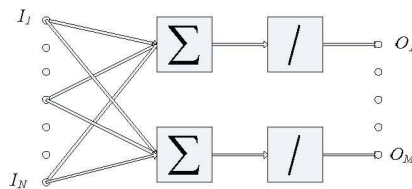

Figure 3. Simple single layer back-propagation neural network used for color restoration

Input images have been split into 6×6 blocks. The size was selected to fit an integer number atomic blocks of each pattern given for evaluation. These blocks have been interpolated by the neural network as a whole.

The number of outputs for one block is $M = C \cdot H_{block} \cdot W_{block}$. Where $C = 3$ is number of colors and $H_{block} = W_{block} = 6$ for every CFA. Width and height have been chosen as least common multiple of dimensions of all selected color filter arrays to ensure that the total number of neurons can be the same for each array tested. Neighbors of size 12×12 have been chosen as inputs of the nets ($N = 12 \times 12 = 144$). The same image was used for learning and evaluation. It was not the intention to create the best possible net for demosaicing using an offline learning scheme, but in contrast to explore the input data in the context of its linear correlation with a visually perceptible output image. In other words there was an attempt to fix the number of biases in the linear demosaicking algorithm and to find which configuration of input samples provides a better resultant image according to subjective perception and MSE for RGB color components.

Neural networks were trained for every image individually. The MSE between the original RGB block and the output vector was used for training. The condition for halting the training process was reaching zero gradient. Since neural networks are linear there is only one minimum in the training function. In practice the minima were found relatively quickly due to linearity on the net and a very small training set (one image). The 1338 images from UCID [3]

have been processed with color restoring neural nets. Restored images have been compared using MSE and NCD. The summary of the comparison is given in Table 2.

TABLE II.   RESULT S OF DEMOSAICING USING NEURAL NETS

	Classic Bayer	3x2 pattern	Diagonal Bayer
Average NCD	0.0790	0.0769	0.0779
Average MSE ratio	1.068	0.961	0.970
Standard deviation of ratio	0.043	0.025	0.045
Maximum MSE ratio	1.159	1.023	1.025
Minimum MSE ratio	1.007	0.920	0.818

IV. CONCLUSION

In this paper three different ways of comparing the performance of color filter arrays have been proposed and demonstrated. Usually when comparing color filter arrays with different spatial configuration the performance depends significantly on demosaicking algorithm selected for specific CFA. Two of proposed methods do not depend on the demosaicking algorithm and are based on averaging of image data. The third method is an attempt to equalize the capabilities of demosaicking algorithms for various configurations of color filter arrays by using back-propagation artificial neural networks with identical configuration. Both statistical error evaluation (MSE) and visual assessment of results have been performed. The applied model of color filter array does not include noise of input data and its color aberration. Also a primitive model of human visual perception has been used. It has been shown that alternatives to classic Bayer CFA can give good performance in some circumstances, and that there may be benefits in using a different pattern.

V. ACKNOWLEDGEMENT

REFERENCES

[1] B. Bayer, "Color Imaging Array," U.S. Patent 3 971 065, 1976.

[2] Gunter Wyszecki, W. S. Stiles, "Color Science: Concepts and Methods, Quantitative Data and Formulae," 2nd Edition, 1982, pp. 514 - 581

[3] G. Schaefer and M. Stich (2004) "UCID - An Uncompressed Color Image Database", *Proc. SPIE, Storage and Retrieval Methods and Applications for Multimedia 2004*, pp. 472-480, San Jose, USA. < http://vision.doc.ntu.ac.uk/datasets/UCID/ucid.html>

[4] K.T. Mullen, "The contrast sensitivity of human color vision to red-green and blue-yellow chromatic grating," *Journal of Physiology*, 359, 1995, pp. 381-400.

[5] Wenmiao Lu, Yap-Peng Tan, "Color Filter Array Demosaicking: New Method and Performance Measures," *IEEE Transactions on Image Processing*, vol. 12, no. 10, Oct 2003

[6] K. N. Platoiotis and A.N. Venetsanopoulos. *Color Image Processing and Applications*. Springer Verlag, 2000.

[7] Malkvar, H. S., He, L., and Cutler, R., 2004, High-quality linear interpolation for demosaicing of Bayer-patterned color images, <research.microsoft.com/users/lhe/papers/icassp04.demosaicing.pdf>